From:	PETERSON Jenn L
To:	Chip Humphrey/R10/USEPA/US@EPA; Burt Shephard/R10/USEPA/US@EPA
Cc:	Eric Blischke/R10/USEPA/US@EPA; ANDERSON Jim M; Kristine Koch/R10/USEPA/US@EPA
Subject:	RE: Scheduling Benthic Follow up meeting 12/2 or 12/3
Date:	11/29/2010 07:50 AM

Chip,

I can do the 2nd before 2:00, but not the 3rd.

Jennifer

-----Original Message-----
From: Humphrey.Chip@epamail.epa.gov [mailto:Humphrey.Chip@epamail.epa.gov]
Sent: Wednesday, November 24, 2010 11:29 AM
To: Shephard.Burt@epamail.epa.gov; Goulet.Joe@epamail.epa.gov; PETERSON Jenn L; POULSEN Mike; Bob Dexter; jay.field@noaa.gov; mesl@shaw.ca; JMalek@parametrix.com; AEbbets@stratusconsulting.com
Cc: Blischke.Eric@epamail.epa.gov; ANDERSON Jim M; Koch.Kristine@epamail.epa.gov
Subject: Scheduling Benthic Follow up meeting 12/2 or 12/3

On the November 4th meeting/conference call with LWG's benthic modeling team we agreed to get the EPA/LWG team back together around December 1st to discuss the action items from the meeting. John Toll has provided a summary of the call (see his two emails below) and action items. We trust that our work is moving forward, and it sounds like the LWG will be ready to meet.

I am checking on our team's availability to meet (or participate in a conference call) on Thursday, December 2nd or Friday, December 3rd.

thanks
Chip

----- Forwarded by Chip Humphrey/R10/USEPA/US on 11/24/2010 11:13 AM -----

RE: 11/4/10 benthic modeling meeting summary

John Toll
to:
John Toll, Eric Blischke, Chip Humphrey
11/18/2010 05:18 PM

Cc:
Jennifer Woronets, James McKenna , Lucinda Tear, Nancy Musgrove, "Lorraine B. Read"

Eric and Chip,

When I met with the benthic analysis team today we came up with two more LWG action items that we've added to our December 1 to do list.

1. Test whether FPM SQVs are affected by the order chemicals go into the model. By the way, the answer's no. Lucinda ran tests and also figured out how DEQ could have concluded that order matters; apparently the FPM software is prone to mistakes if you try playing around with order.
2. Peer review the statistics write up that you gave us at the September 29 benthic check-in.

As you saw, Jay wrote today asking for the final hit classification table and commenting on what we called a NOAA QC error at station G643. We're looking into G643 and will send the finalized table tomorrow.

Again, please let me know if you see anything in my meeting summary that you think is inaccurate, or if you think we missed any decisions or actions. We're trying to avoid any miscommunications that could set back the schedule.

Thanks,

John

From: John Toll
Sent: Wednesday, November 17, 2010 6:46 PM
To: Eric Blischke (Blischke.Eric@epamail.epa.gov); 'Chip Humphrey (humphrey.chip@epa.gov)'

Cc: Jennifer Woronets; 'James McKenna'
Subject: 11/4/10 benthic modeling meeting summary

Eric and Chip,

Following is a summary I wrote of our 11/4/10 benthic modeling meeting. I'm sorry I wasn't able to send it to you sooner. We think that this summary accurately captures all the decisions and actions that came out of the meeting. Please let me know if you see anything that you think is inaccurate, or if you think we missed any decisions or actions. Thanks,

John

Attendees:

Don MacDonald, Bob Dexter, Eric Blischke, Mark Lewis, Jay Field, Keith Pine, Jennifer Peterson, Mike Poulson, Mike Carney (Stratus), Nancy Musgrove, Chip Humphrey, Kevin Parrett, Rick Applegate, Lorraine Read, Bob Wyatt, Joe Goulet, Burt Shephard, John Toll

Summary:

The LWG's benthic modeling team met with the agency team from 8:30-11:00 AM on 11/4/10. The purpose of the meeting was to meet the first of a series of three sets of milestones established by Eric Blischke in a 10/27/10 meeting (attended by Eric Blischke, Jim McKenna, John Toll, Nancy Musgrove, Chip Humphrey and Burt Shephard). The first four major headings below correspond to the first set of milestones. All four milestones were met. Decisions and actions are summarized below. The next milestone is on or around 12/1/10. The actions identified below are to be completed by then.

1. Hit classification reconciliation – The LWG identified all hit classification differences and the reason for each. Of the 1,172 bioassay results there were 27 hit classification differences. 25 were due to differences in at which step in the process the LWG and the agency team rounded to two significant figures. The other two were QC error (one by the LWG, one by the agency team). Both sides rounding rules were reasonable and there's no definitive technical argument for one or the other. Using the agency team's rounding rule won't add to the amount of work we have to do. The effect of switching to the agency team's rounding rule is not detrimental to the LWG (4 stations decrease from L1 to L0, 4 decrease from L2 to L1, 9 decrease from L3 to L2, six increase from L0 to L1 and two increase from L1 to L2; see the map that Jen forwarded to Exec on the evening of 11/3 under the subject line "FW: information to be transmitted to EPA for tomorrow's benthic meeting").
Decisions and Actions
   - The LWG will update its hit calculations table to reflect the agency team's rounding method and correct the one LWG QC error and send the updated table to Eric. This will be the final master table of bioassay hit classifications moving forward.

2. Logistic Regression Model – We discussed the LWG's bullet list of requests for LRM documentation (memo from Toll to Blischke, provided to Exec as part of what Jen forwarded on the evening of 11/3 under the subject line "FW: information to be transmitted to EPA for tomorrow's benthic meeting") and how the LRM will be used to predict L3 hits.
Decisions and Actions
   - The agency team will provide sufficient documentation of its September 2010 LRM for a qualified independent modeler to be able to reproduce it. This will include the information requested in the LWG's memo. Sufficient documentation for a qualified independent modeler to be able to reproduce the LRM is due in rough form by Dec 1 with final documentation to be attached to the revised draft BERA.
   - All models including the agency team's September 2010 LRM will pre-process the sediment chemistry data used for model development in accordance with BERA rules (e.g., data quality and summing rules).
   - All models including the agency team's September 2010 LRM will use the reconciled hit classifications as documented in the final master table of bioassay hit classifications.
   - The agency team will report reliability statistics on the LRM when it meets again with the agency team on ~12/1. This will include at a minimum all the reliability statistics provided by EPA with its comments on the draft BERA. Documentation (sufficient for an independent modeler to reproduce results) is required for any other reliability statistics that are run.
   - The LRM will be used to predict L3 hits by setting a higher Pmax threshold. The agency team will recommend a Pmax threshold based on a documented analysis of model reliability statistics.

3. Floating Percentile Model – We discussed verbal requests from the agency team for FPM documentation. We also discussed modeling assumptions, specifically, using individual endpoints versus a pooled endpoint, methods for

eliminating chemicals from the model (detection frequency, use of parametric versus non-parametric ANOVA), varying false negative rate, summing rules and allowance for "subjective improvements" based on the modelers' BPJ.

Decisions and Actions

- The decision of whether to use individual or pooled endpoints is at the LWG's discretion. The LWG's modelers indicated a preference for using individual endpoint FPMs. EPA agreed that it is not necessary to carry forward both individual and pooled endpoint FPMs.
- EPA agreed that chemicals may be eliminated based on detection frequency but asked for clearer explanation of how and why chemicals are eliminated.
- The LWG has in the past varied the false negative rate in the FPM at 5% increments. EPA asked for clearer documentation of this and a full set of reliability statistics on all runs. EPA indicated that it was interested in seeing results for false negative rates ranging from 5-25% (in 5% increments).
- EPA requested clearer documentation of the use of parametric versus non-parametric ANOVA to screen FPM chemicals.
- EPA requested equivalent documentation of the L2 and L3 FPMs.
- The LWG will provide sufficient documentation of the individual endpoint FPMs for a qualified independent modeler to be able to reproduce it. Sufficient documentation for a qualified independent modeler to be able to reproduce the FPMs is due in rough form by Dec 1 with final documentation to be attached to the revised draft BERA.
- All models including the individual endpoint FPMs will pre-process the sediment chemistry data used for model development in accordance with BERA rules (e.g., data quality and summing rules).
- All models including the individual endpoint FPMs will use the reconciled hit classifications as documented in the final master table of bioassay hit classifications.
- The LWG will report reliability statistics on the individual endpoint FPMs when it meets again with the agency team on ~12/1. This will include at a minimum all the reliability statistics provided by EPA with its comments on the draft BERA. Documentation (sufficient for an independent modeler to reproduce results) is required for any other reliability statistics that are run.
- EPA stated that subjective improvements to the FPMs will be allowed at the BPJ of the modelers, but that any such improvements must be documented sufficiently for a qualified independent modeler to be able to reproduce the FPMs.

4. Additional Models – We discussed the mean quotient method that was used in the comprehensive benthic approach for alternatives screening. We also discussed the use of generic SQG sets.

Decisions and Actions

- The LWG's modelers agreed to provide documentation of the mean quotient method that was used in the comprehensive benthic approach for alternatives screening.
- EPA agreed that the only generic SQG sets that should be carried forward are PECs and PELs.
- EPA asked the LWG to run reliability statistics on PECs and PELs as individual SQG sets and on PEC and PEL mean quotients.

5. Other –

- Eric Blischke stated that the ultimate purpose of benthic toxicity modeling is to provide information to risk managers for setting cleanup levels.
- There was a fairly extensive discussion of using randomization tests to evaluate whether the benthic toxicity models are performing better than random chance performance. The agency team generally agreed that this is valuable in principle but raised concerns about how much work would be involved. Lorraine Read described how she's done it in the past on the LWG's version of the LRM and that it's not time-consuming (she explained how she set up the randomization run inputs in a batch file and ran the analysis overnight). The agency team wasn't willing to commit to running randomization tests on its LRM but didn't object to the LWG's introducing them as a model evaluation tool.
- The LWG described (again) it's likelihood analysis of bioassay response levels and explained why it is pertinent, in consort with reliability statistics, for selecting LRM L2 and L3 Pmax values. Eric Blischke noted that the agency team would need to see documentation before it could be considered.

John Toll, Ph.D.
Partner, Windward Environmental LLC

200 W. Mercer Street, Suite 401 |
Seattle, WA 98119
206.812.5433 (o) | 206.913.3292 (c)
johnt@windwardenv.com |
www.windwardenv.com